SOME ELEMENTARY CONCEPTS OF MATRIX ALGEBRA

AND ITS APPLICATIONS

F. E. Walker

Department of Agricultural Economics
The Ohio State University

Departmental Series A. E. 351

September, 1963

Some Elementary Concepts of Matrix Algebra
And Its Applications

References:

1. Klein, Lawerence R., A Textbook of Econometrics, Row, Peterson and Company, 1953, pp. 324 - 345.

2. Tintner, G., Econometrics, John Wiley, 1952, appendix.

3. Friedman, Joan and Richard J. Foote, "Computational Methods for Handling Systems of Simultaneous Equations", U. S. Department of Agriculture, Agriculture Handbook No. 94, November, 1955, pp. 21-27.

4. Graybill, Franklin, An Introduction to Linear Statistical Models, McGraw-Hill, 1961. This book contains a section on the rules for differentiating matrices.

5. Dorfman, Robert, Paul A. Samuelson, and Robert M. Solow, Linear Programming and Economic Analysis, McGraw-Hill, 1958, pp. 470-506.

6. Allen, R.G.D., Mathematical Economics, St. Martin's Press, New York, 1957.

7. Kemeny, J. G., J. L. Snell, and G. L. Thompson, Finite Mathematics, Prentice-Hall, 1957.

For a more complete discussion of linear transformations, vectors, and matrices, see:

1. Hohn, Franz, Elementary Matrix Algebra, Macmillan, 1958.

2. Perlis, S., Theory of Matrices, Addison-Wesley Publishing Company, Cambridge, Mass., 1952.

3. Aitken, A. C., Determinants and Matrices, 8th edition, Interscience, New York, 1954.

I. Definition

A matrix is a rectangular array of numbers. The elements (numbers) making up a matrix represent some entity (a coefficient, a variable, etc.). For example, given the system of two equations in two unknowns,

$$\begin{aligned} 1\,x + 2\,y &= 3 \\ -3\,x + 4\,y &= 7, \end{aligned} \tag{1.1}$$

matrix notation may be used to write the system in the more compact form,

$$A\,X = B, \tag{1.2}$$

if the matrices (A, X, and B) are appropriately defined. Furthermore, we can perform operations on these matrices in order to solve for the unknowns, x and y. These matrix operations correspond to the scalar (one element or number) operations used to solve for x and y in (1.1). The appropriate definition of A, X, and B of (1.2) in order to be equivalent to (1.1) is

A = the detached coefficients of x and y, such that the coefficients of the first equation make up the first row, the coefficients of equation 2 of (1.1) compose row 2, the coefficients of x and y are the first and second columns respectively.

$$A = \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix}$$

$X = \begin{bmatrix} x \\ y \end{bmatrix}$ = a matrix of one column and two rows whose elements are the unknowns.

$B = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$ = a one column matrix of the right-hand terms (constants) of each of the two equations.

System (1.1) can then be written as (1.2) or specifically as

$$\begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 3 \\ 7 \end{bmatrix} \tag{1.3}$$

However note that all operations can be performed on (1.2) without knowing the specific form (1.3).

It will be assumed in this discussion that the elements of each matrix are real numbers or variables which can take on the values of real numbers only. Matrix algebra methods apply to complex number elements with some modifications of a few theorems.

The term a_{ij} will be used to designate the element in the i'th row and j'th column of matrix A. The element in row 1 and column two of A is designated a_{12} and in the example is 2.

A matrix is said to be of order mxn if it has m rows and n columns. A is of order 2x2, X is 2x1, and B is 2x1.

The principal diagonal of a matrix (applicable only to a matrix which has an equal number of rows and columns) is the set of elements composing the diagonal line beginning in the upper left corner and extending to the lower right hand corner of the matrix.

II. Equality of matrices. Two matrices, A and B, are defined to be equal if and only if (iff) $a_{ij} = b_{ij}$. This implies that the order of A is the same as the order of B.

III. Addition and subtraction of matrices.

In order to add or subtract two or more matrices they must be conformable for addition or subtraction, i.e. their orders must be the same.

To add (subtract), corresponding elements of the matrices, A and B, are added (subtracted). To form A + B = C, find $a_{ij} + b_{ij} = c_{ij}$. To form A - B = D, find $a_{ij} - b_{ij} = d_{ij}$.

Example:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$$B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} = \begin{bmatrix} 2 & 0 & 8 \\ 3 & 1 & 4 \end{bmatrix}$$

$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \end{bmatrix}$$

$$= \begin{bmatrix} 3 & 2 & 11 \\ 7 & 6 & 10 \end{bmatrix}$$

The commutative (a + b = b + a) law and the associative a + (b + c) = (a + b) + c laws of scalar algebra hold for matrix algebra, i.e.

$$A + B = B + A$$

and $$A + (B+C) = (A+B) + C$$

A matrix may be multiplied by a scalar (α).

$$\alpha \cdot A = \begin{bmatrix} \alpha a_{ij} \end{bmatrix} = A \cdot \alpha$$

If $\alpha = 3$ and
$$A = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix},$$
$$\alpha \cdot A = \begin{bmatrix} 3 & 9 \\ 6 & 12 \end{bmatrix}$$

IV. Matrix multiplication

If (1.2) is to be equivalent to (1.1) then the product $A \cdot X$ of (1.2) must be formed in a manner which will form the left hand side of the system (1.1). This process is matrix multiplication.

Matrix multiplication (AB) is performed by multiplying the elements of each row of the first matrix (A) by the corresponding elements of each column of the second matrix (B) and forming the sum of these products, one sum for each possible combination of a row and column. These sums form the elements of the product matrix.

A and B are conformable for multiplication (AB =C) if matrix A has as many columns as there are rows in B. The product matrix (C) will contain the same number of rows as A and the same number of columns as B. Thus if A is of order mxn and B of order nxp, A and B are conformable for multiplication and C is of order mxp. The product matrix, AB, has a number of rows equal to the number of rows in the first matrix (A) of the product and a number of columns equal to the number of columns of the last (second) matrix (B).

Example:
$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} ; \quad B = \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & c_{23} \end{bmatrix}$$

and
$$C = \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \end{bmatrix}, \text{ where}$$

$$c_{11} = a_{11} b_{11} + a_{12} b_{21}$$
$$c_{12} = a_{11} b_{12} + a_{12} b_{22}$$
$$c_{13} = a_{11} b_{13} + a_{12} b_{23}$$
$$c_{21} = a_{21} b_{11} + a_{22} b_{21}$$
$$c_{22} = a_{21} b_{12} + a_{22} b_{22}$$
$$c_{23} = a_{21} b_{13} + a_{22} b_{23}$$

If

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 4 & 3 & 2 \\ 1 & 0 & 5 \end{bmatrix},$$

then

$$C = AB = \begin{bmatrix} 6 & 3 & 12 \\ 16 & 9 & 26 \end{bmatrix}$$

Care must be taken in the arrangement of matrices for multiplication. A and B are conformable for forming the product AB but not conformable for forming BA. Pre-multiplication of B by A is the term given to forming the product AB. This could also be termed post-multiplication of A by B. Ordinarily the commutative law of scalar algebra (ab = ba) will not hold for matrix algebra, i. e.

$$AB \neq BA$$

Therefore, a distinction must be made between pre- and post-multiplication of A by B.

Example:

$$A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} ; \quad B = \begin{bmatrix} 1 & 4 \\ -1 & 1 \end{bmatrix}$$

$$AB = \begin{bmatrix} 3 & 7 \\ -3 & -2 \end{bmatrix} \neq \begin{bmatrix} -2 & 7 \\ -3 & 3 \end{bmatrix} = BA$$

In scalar algebra if ab =0 , then either a or b is equal to zero. It is not true in matrix algebra that if AB = O, then either A or B must be the zero matrix. O is a matrix whose elements are all zero.

Example:

$$AB = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 4 & 9 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

yet neither A nor B is the zero matrix. Also it is possible to have AB = AC without having B = C.

Example:

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 1 & 1 & 0 \\ -1 & 4 & 0 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 2 & 4 & 2 \end{bmatrix}$$

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 1 & -1 \\ 1 & 5 & 7 \end{bmatrix}$$

$$AB = \begin{bmatrix} 3 & 4 & 1 \\ 2 & 3 & 2 \\ 3 & 2 & -7 \end{bmatrix} = AC$$

yet $B \neq C$.

In summary, there are three fundamental properties of matrix multiplication which are different from those of scalar multiplication.

1. The rule AB = BA does not hold generally.
2. From AB = O, one cannot conclude that either A or B is a zero matrix.
3. From AB = AC, one cannot conclude that B = C.

The associative law and the distributive law do hold for matrix algebra, i.e.

$$(AB)\,C = A\,(BC)$$
$$A\,(B + C) = AB + AC.$$

V. Linear equation systems in matrix form.

The definition of the matrix multiplication process allowed us to write (1.1) as (1.2). Any system of equations in which there are m equations in n unknowns can be written in matrix form for computational purposes. This general system written in scalar form is:

$$\begin{array}{l} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 \dots + a_{mn}x_n = b_m \end{array} \qquad (5.1)$$

Note that a_{ij} is the coefficient of the j'th variable in the i'th equation. (5.1) in matrix form is

$$AX = B$$

where

(5.1.1)
$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2n} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix},$$

X is the column vector (one column matrix)

(5.1.2)
$$\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix},$$

and B is the column vector

(5.1.3)
$$\begin{bmatrix} b_1 \\ b_2 \\ \cdot \\ \cdot \\ \cdot \\ b_m \end{bmatrix}$$

Note that X must have n rows to be conformable for pre-multiplication by A.

An example of the use of matrices in solution of a problem is the use of the Doolittle solution in finding the least squares fit of a linear equation to a set of observed equations. The normal equations of least squares form a system of equations which can be written in matrix form. Given the following equation,

(5.2) $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3$

to be fit to a set of data, the normal equations are formed.

$$N_a + b_1 \sum X_1 + b_2 \sum X_2 + b_3 \sum X_3 = \sum Y$$

$$a \sum X_1 + b_1 \sum X_1^2 + b_2 \sum X_1 X_2 + b_3 \sum X_1 X_3 = \sum X_1 Y$$

(5.3) $$a \sum X_2 + b_1 \sum X_1 X_2 + b_2 \sum X_2^2 + b_3 \sum X_2 X_3 = \sum X_2 Y$$

$$a \sum X_3 + b_1 \sum X_1 X_3 + b_2 \sum X_2 X_3 + b_3 \sum X_3^2 = \sum X_3 Y$$

N is the number of observations; $\sum X_1$, $\sum X_2$, $\sum X_3$, and $\sum Y$ are the sums of the observed values of the 4 variables; and $\sum X_1^2$, $\sum X_2^2$, $\sum X_3^2$, $\sum X_1X_2$, $\sum X_1X_3$, $\sum X_1 Y$, $\sum X_2X_3$, $\sum X_2 Y$, and $\sum X_3 Y$ are the sums of squares and cross products of the 3 independent variables and sums of the cross products of the dependent variable (Y) with each independent variable.

(5.3) may be written

(5.4) $$XB = Y$$

where

(5.5) $$X = \begin{bmatrix} N & \sum X_1 & \sum X_2 & \sum X_3 \\ \sum X_1 & \sum X_1^2 & \sum X_1X_2 & \sum X_1X_3 \\ \sum X_2 & \sum X_1X_2 & \sum X_2^2 & \sum X_2X_3 \\ \sum X_3 & \sum X_1X_3 & \sum X_2X_3 & \sum X_3^2 \end{bmatrix}$$

(5.6) $$B = \begin{bmatrix} a \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad ; \quad Y = \begin{bmatrix} \sum Y \\ \sum X_1 Y \\ \sum X_2 Y \\ \sum X_3 Y \end{bmatrix}$$

If the variables are measured as deviations from their respective means $(X_1 - \overline{X}_1)$, then the system of normal equations contains 3 equations in 3 unknowns, b_1, b_2, and b_3. The matrix form of the system would still be (5.4) with

(5.7) $$X = \begin{bmatrix} \sum x_1^2 & \sum x_1x_2 & \sum x_1x_3 \\ \sum x_1x_2 & \sum x_2^2 & \sum x_2x_3 \\ \sum x_1x_3 & \sum x_2x_3 & \sum x_3^2 \end{bmatrix}$$

(5.8) $$B = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \quad ; \quad Y = \begin{bmatrix} \sum x_1 y \\ \sum x_2 y \\ \sum x_3 y \end{bmatrix}$$

Lower case x and y indicate that the variables have been measured as deviations from means; therefore, $\sum y$, $\sum x_1$, $\sum x_2$, and $\sum x_3$ are all zero. The Doolittle method is equivalent to solving XB = Y for B.

VI. Definitions of special matrices.

Transpose of a matrix (A^T or A'). A' is formed from A by interchanging the rows and columns of A, i.e. row 1 of A becomes column 1 of A', the transpose, and vice versa. Thus the transpose of A_{mxn} (a matrix of order mxn) is A'_{nxm}.

Example:

$$A_{2x3} = \begin{bmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{bmatrix} \underset{2x3}{=} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$A'_{3x2} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix}$$

The transpose of the transpose of A is equal to A, i.e.

$$(A')' = A$$

The transpose of the product of two matrices (AB) is the product of the transpose matrices taken in reverse order, ie.

$$(AB)' = B'A'$$
or
$$(ABC)' = C'B'A'$$

Symmetric matrix. A symmetric matrix is a square matrix (m=n) in which $a_{ij} = a_{ji}$, e.g. $a_{12} = a_{21}$ or $a_{41} = a_{14}$. Thus for every number below the principal diagonal, there is a like number above the principal diagonal. A matrix, A, is symmetric iff $A' = A$.

Example:

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

$$A' = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}$$

Note that $a_{12} = a_{21}$, $a_{11} = a_{11}$, $a_{22} = a_{22}$

In multiple regression analysis, the matrix of sums of squares and cross products is always symmetric.

Scalar matrix. A scalar matrix is a square matrix in which the elements of the principal diagonal are all equal. All non-diagonal elements are zero.

Example: $\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix}$ or $\begin{bmatrix} \alpha & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & \alpha \end{bmatrix}$

where α can be any constant. Multiplying a matrix by the scalar matrix (with α as the constant) is equivalent to multiplying by a constant in scalar algebra.

Identity matrix. If in the scalar matrix $\alpha = 1$, the matrix is called the identity matrix (or unit matrix), I. The identity matrix of order 3 is

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplication of any matrix (A) by an identity matrix (I) has the effect of reproducing the original matrix, ie.

$$IA = A$$

This operation is analogous to multiplication by 1 in scalar algebra.

Example: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 6 & 5 & 4 \\ 7 & 8 & 9 \end{bmatrix}$

Note that I is a symmetric matrix.

VI. Determinants

Determinants are associated with square matrices only and are a value which represents the "magnitude" of this square matrix. Computation of a determinant is essential in the operation of "dividing" with matrices (matrix inversion).

The determinant of the matrix A is written $|A|$ or det A.

If A is the 1 - element matrix $[a_{11}]$, the det $A = a_{11}$. For the 2x2 matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

$\det A = a_{11}\, a_{22} - a_{21}\, a_{12}$.

Finding the determinant of higher order matrices becomes computationally more difficult. If A is the matrix

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

its determinant can be found by writing the matrix with the first two columns appended as two additional columns on the right of the matrix.

$$\left[\begin{array}{ccc|cc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{array}\right]$$

Draw arrows diagonally through each possible set of three elements, such that the arrows are pointing either up to the right or down to the right. Form the product of the elements along each of the six arrows. The determinant of A is found as the sum of the products of the arrows pointing down minus the sum of the products for the arrows pointing up. This procedure is appropriate only for matrix of order 3x3.

$$\text{Thus det } A = a_{11}\, a_{22}\, a_{33} + a_{12}\, a_{23}\, a_{31} + a_{13}\, a_{21}\, a_{32} - a_{31}\, a_{22} a_{13} - a_{32}\, a_{23}\, a_{11} - a_{33}\, a_{21}\, a_{12}$$

Example:

(6.1) $$A = \begin{vmatrix} -6 & 11 & -8 \\ 2 & -3 & 0 \\ 2 & -5 & 4 \end{vmatrix}$$

$$\begin{aligned} \det A &= (-6)\,(-3)\,(4) + (11)\,(0)\,(2) + (-8)\,(2)\,(-5) - (2)\,(-3)\,(-8) - (-5)\,(0)\,(-6) - (4)\,(2)\,(11) \\ &= 16 \end{aligned}$$

One seldom will be required to find the determinants of higher order matrices. The procedure is complex and the results are unnecessary in matrix operations. Only in matrix operations with small matrices is it feasible, with respect to time, to find the determinants.

A determinant of a matrix of any order can be evaluated by the method of expanding along any row or any column. A short discussion of this method will aid in the understanding of the process of matrix inversion. A few new terms used in this method need to be defined. A_{ij} is the minor of a_{ij}, the element of the matrix A. The minor of a_{ij} is formed by deleting from A the i'th row and the j'th column. The portion of A remaining is A_{ij}.

Example:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

The minor of a_{21} ($i = 2,\ j = 1$) is formed by removing the second row and the first column. Thus

$$A_{21} = \begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} = \text{the minor of } a_{21}$$

The cofactor of a_{ij}, an element of the square matrix A, is the signed determinant of A_{ij}, ie. the cofactor of a_{ij}, cof a_{ij}, is

$$(-1)^{i+j} \det A_{ij}$$

The sign of the cofactor is dependent upon the row and column subscripts of the particular element of matrix A. In the preceding example, cof a_{21} is

$$(-1)^{2+1} \det A_{21} \text{ or } - \begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix}$$

$$\text{cof } a_{21} = -a_{12}\, a_{33} + a_{32}\, a_{13}$$

To find the determinant of A one might "expand" along any row. In this process we find the determinant by summing the products of each element of any row times the cofactor of the element. det A of the matrix A (above) can be found by expanding along row one.

$$\det A = a_{11} \cdot \text{cof } a_{11} + a_{12} \cdot \text{cof } a_{12} + a_{13} \cdot \text{cof } a_{13}$$

We could also find det A by expanding along row two or three or any one of the three columns, ie. det A is also equal to, by expanding all column three

$$a_{13} \cdot \text{cof } a_{13} + a_{23} \cdot \text{cof } a_{23} + a_{33} \cdot \text{cof } a_{33}$$

Example: The determinant of (6.1) when expanded along row two is

$$\det A = a_{21} \cdot \text{cof } a_{21} + a_{22} \cdot \text{cof } a_{22} + a_{23} \cdot \text{cof } a_{23}$$

$$= 2\,(-1)^{2+1} \begin{vmatrix} 11 & -8 \\ -5 & 4 \end{vmatrix}$$

$$+ (-3)(-1)^{2+2} \begin{vmatrix} -6 & -8 \\ 2 & 4 \end{vmatrix}$$

$$+ 0 (-1)^{2+3} \begin{vmatrix} -6 & 11 \\ 2 & -5 \end{vmatrix}$$

$$= -2 (4) - 3 (-8) - 0 (8)$$

$$= 16$$

The general rule for evaluating a determinant of matrix A_{nxn} by expanding along row i is

$$\det A = \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \cdot \det A_{ij}$$

VII. Matrix inversion.

The matrix operation analogous to division by a number in scalar algebra is matrix inversion. The inverse of a matrix is applicable only to square matrices. The inverse of A, written A^{-1}, is the matrix which when multiplied by A will give the identity matrix, I. In scalar algebra a number, a, times its reciprocal $\frac{1}{a}$, is equal to 1. Similarly the matrix, A, times its inverse (reciprocal), A^{-1}, is equal to I, which in matrix algebra is analogous to 1 in scalar algebra. Thus, one has

(7.1) $$A^{-1} \cdot A = A \cdot A^{-1} = I$$

Note that the orders of A, A^{-1}, and I must all be the same. Only <u>square</u> matrices have inverses and then only if $\det A \neq 0$.

Using the matrix inversion operation, the solution to (1.2) can be written entirely in matrix form.

(7.2) $$AX = B$$

Pre-multiply each side by the inverse of the square matrix A.

(7.3) $$A^{-1} AX = A^{-1} B.$$

$A^{-1} A = I$ and the multiplication of a matrix by I reproduces the matrix. Therefore

(7.4) $$A^{-1} AX = IX = X = A^{-1} B,$$

and x and y, the elements of X, are equal to the elements of the product matrix, $A^{-1} B$.

Example:

(1.1)

$$x + 2y = 3$$
$$-3x + 4y = 7$$

$$[A] = \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix} \quad ; \quad B = \begin{bmatrix} 3 \\ 7 \end{bmatrix} \quad ; \quad X = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\det A = 1 \cdot 4 - (-3)2 = \underline{10}$$

$$A^{-1} = \begin{bmatrix} .4 & -.2 \\ .3 & .1 \end{bmatrix}$$

Substituting A^{-1} and B in (7.4),

$$X = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} .4 & -.2 \\ .3 & .1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{bmatrix} -.2 \\ 1.6 \end{bmatrix}$$

To see that the computed A^{-1} is really an inverse of A, form

$$A^{-1} \cdot A = \begin{bmatrix} .4 & -.2 \\ .3 & .1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I;$$

therefore the computed A^{-1} is an inverse of A.

The normal equations of a single equation regression model are given by (5.4) with the associated matrices defined by (5.7) and (5.8). The Doolittle method is essentially a method of computing.

$$B = X^{-1}Y$$

for from (5.4)

$$XB = Y$$
$$X^{-1}XB = X^{-1}Y$$
$$X^{-1}XB = IB = B = X^{-1}Y.$$

Thus $X^{-1}Y$ gives the regression coefficients b_1, b_2 and b_3 of (5.2).

Computing the inverse of a matrix of order 1 or 2 is simple and can be written in a single formula. Inverse matrices of order 3 can be readily computed but for higher order matrices, standard procedures are best adapted for the computations. See Freidman (3).

Inverse of a 1 element matrix.

$$A_{1x1} = \begin{bmatrix} a_{11} \end{bmatrix}$$
$$\det A = a_{11}$$
$$A^{-1} = \begin{bmatrix} \frac{1}{a_{11}} \end{bmatrix}$$

Inverse of a 2x2 matrix.

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\det A = a_{11} a_{22} - a_{12} a_{21}$$

$$A^{-1} = \frac{1}{\det A}\begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} = \begin{bmatrix} \frac{a_{22}}{\det A} & \frac{-a_{12}}{\det A} \\ \frac{-a_{21}}{\det A} & \frac{a_{11}}{\det A} \end{bmatrix}$$

Inverse of a 3x3 matrix.

The inverse of a matrix, A, may be defined as the transposed matrix of cofactors of A divided by det A. This definition suggests the approach to the inversion of the 3x3 matrix A.

1. Step 1

   Compute det A.

2. Step 2

   Form the matrix of cofactors.

$$\begin{bmatrix} \text{cof } a_{11} & \text{cof } a_{12} & \text{cof } a_{13} \\ \text{cof } a_{21} & \text{cof } a_{22} & \text{cof } a_{23} \\ \text{cof } a_{31} & \text{cof } a_{32} & \text{cof } a_{33} \end{bmatrix}$$

3. Step 3.

   Transpose the matrix of cofactors.

$$\begin{bmatrix} \text{cof } a_{11} & \text{cof } a_{21} & \text{cof } a_{31} \\ \text{cof } a_{12} & \text{cof } a_{22} & \text{cof } a_{32} \\ \text{cof } a_{13} & \text{cof } a_{23} & \text{cof } a_{33} \end{bmatrix}$$

4. Step 4.

   Divide each element of the matrix of cofactors by det A. This result is the 3x3 matrix, A^{-1}.

$$A^{-1} = \begin{bmatrix} a^{11} & a^{12} & a^{13} \\ a^{21} & a^{22} & a^{23} \\ a^{31} & a^{32} & a^{33} \end{bmatrix}$$

a^{ij} indicates the element of an inverse matrix in the i'th row and j'th column.

Example:

Compute the inverse of

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Step 1.

Computation of det A is explained in section VI and will not be explained in detail here.

Step II.

The matrix of cofactors is:

$$\begin{bmatrix} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ -\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix} \\ \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} & -\begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix} & \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \end{bmatrix}$$

Step III.

The transposed matrix of cofactors with the determinants evaluated is:

$$\begin{bmatrix} a_{22}\ a_{33} - a_{32}\ a_{23} & a_{32}\ a_{13} - a_{12}\ a_{33} & a_{12}\ a_{23} - a_{13}\ a_{22} \\ a_{31}\ a_{23} - a_{21}\ a_{33} & a_{11}\ a_{33} - a_{13}\ a_{31} & a_{21}\ a_{13} - a_{11}\ a_{23} \\ a_{21}\ a_{32} - a_{31}\ a_{22} & a_{31}\ a_{12} - a_{11}\ a_{32} & a_{11}\ a_{22} - a_{21}\ a_{12} \end{bmatrix}$$

Step IV.

The inverse of A is the matrix of step 3 divided by det A.

$$A^{-1} = \begin{bmatrix} \dfrac{a_{22}\ a_{33} - a_{32}\ a_{23}}{\det A} & \dfrac{a_{32}\ a_{13} - a_{12}\ a_{33}}{\det A} & \dfrac{a_{12}\ a_{23} - a_{13}\ a_{22}}{\det A} \\ \dfrac{a_{31}\ a_{23} - a_{21}\ a_{33}}{\det A} & \dfrac{a_{11}\ a_{33} - a_{13}\ a_{31}}{\det A} & \dfrac{a_{21}\ a_{13} - a_{11}\ a_{23}}{\det A} \\ \dfrac{a_{21}\ a_{32} - a_{31}\ a_{22}}{\det A} & \dfrac{a_{31}\ a_{12} - a_{11}\ a_{32}}{\det A} & \dfrac{a_{11}\ a_{22} - a_{21}\ a_{12}}{\det A} \end{bmatrix}$$

Example:

Find the inverse of

$$A = \begin{bmatrix} 8 & 4 & 2 \\ 2 & 8 & 4 \\ 1 & 2 & 8 \end{bmatrix}$$

Step 1:

det A = 392

Step 2: Matrix of cofactors

$$\begin{bmatrix} \begin{vmatrix} 8 & 4 \\ 2 & 8 \end{vmatrix} & -\begin{vmatrix} 2 & 4 \\ 1 & 8 \end{vmatrix} & \begin{vmatrix} 2 & 8 \\ 1 & 2 \end{vmatrix} \\ -\begin{vmatrix} 4 & 2 \\ 2 & 8 \end{vmatrix} & \begin{vmatrix} 8 & 2 \\ 1 & 8 \end{vmatrix} & -\begin{vmatrix} 8 & 4 \\ 1 & 2 \end{vmatrix} \\ \begin{vmatrix} 4 & 2 \\ 8 & 4 \end{vmatrix} & -\begin{vmatrix} 8 & 2 \\ 2 & 4 \end{vmatrix} & \begin{vmatrix} 8 & 4 \\ 2 & 8 \end{vmatrix} \end{bmatrix}$$

or, evaluating the determinants

$$\begin{bmatrix} 56 & -12 & -4 \\ -28 & 62 & -12 \\ 0 & -28 & 56 \end{bmatrix}$$

Step 3: Transposed matrix of cofactors.

$$\begin{bmatrix} 56 & -28 & 0 \\ -12 & 62 & -28 \\ -4 & -12 & 56 \end{bmatrix}$$

Step 4: Divide by det A = 392 to obtain inverse.

$$A^{-1} = \begin{bmatrix} \frac{56}{392} & -\frac{28}{392} & \frac{0}{392} \\ -\frac{12}{392} & \frac{62}{392} & -\frac{28}{392} \\ -\frac{4}{392} & -\frac{12}{392} & \frac{56}{392} \end{bmatrix} = \begin{bmatrix} \frac{1}{7} & -\frac{1}{14} & 0 \\ -\frac{3}{98} & \frac{31}{196} & -\frac{1}{14} \\ -\frac{1}{98} & -\frac{3}{98} & \frac{1}{7} \end{bmatrix}$$

As a check, compute AA^{-1}.

$$\begin{bmatrix} 8 & 4 & 2 \\ 2 & 8 & 4 \\ 1 & 2 & 8 \end{bmatrix} \begin{bmatrix} \frac{1}{7} & -\frac{1}{14} & 0 \\ -\frac{3}{98} & \frac{31}{196} & -\frac{1}{14} \\ -\frac{1}{98} & -\frac{3}{98} & \frac{1}{7} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Since $A \cdot A^{-1} = I$, the computed A^{-1} is the inverse of A.

In statistics, the moment matrix (matrix of sums of squares and cross products in least squares analysis) is symmetric, i.e. $a_{ij} = a_{ji}$ for all i and j. The inverse of a symmetric matrix is always symmetric. Thus the elements below the principal diagonal need not be computed since they have an equal counterpart above the diagonal.

If in the system of equations (1.1), we form the matrix, A_j, which is the A matrix (coefficient matrix) with the j'th column replaced by the matrix (vector) B, then by Cramer's Rule,

$$x = \frac{\det A_1}{\det A}, \text{ and}$$

$$y = \frac{\det A_2}{\det A}.$$

For the general system of equations (5.1) written in matrix form

$$AX = B$$

where A, X, and B are defined by (5.1.1 - 5.1.3) we find that

$$x_j = \frac{\det A_j}{\det A} \quad \text{for } j = 1, 2, \cdots, n.$$

VIII. Rank of a matrix.

A matrix is said to be of rank r iff it has at least one non-zero determinant of order r, but has no non-zero determinant of order greater than r. The matrix

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & -2 \end{bmatrix}$$

is of rank two for there are no determinants of order 3 but 3 submatrices of order 2. These are $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$, $\begin{bmatrix} 2 & -1 \\ 0 & -2 \end{bmatrix}$, and $\begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}$.

The determinant of each of these is non-zero; therefore A is of rank 2.

The concept of rank is used in connection with the solution (s) to systems of equations. If the n equations of a system of n equations in n unknowns are not independent (i.e. if the rank of the coefficient matrix is not n) one cannot find a unique solution for the n unknowns.

Example:

Find the rank of

$$A = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & -2 \\ 2 & 4 & -3 \end{bmatrix}$$

det A = O, therefore rank is less than 3. The 2x2 submatrices of A are

$$\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 0 & -2 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}, \begin{bmatrix} 2 & 1 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 2 & -1 \\ 2 & -3 \end{bmatrix}, \begin{bmatrix} 1 & -1 \\ 4 & -3 \end{bmatrix},$$

$$\begin{bmatrix} 0 & 3 \\ 2 & 4 \end{bmatrix}, \begin{bmatrix} 0 & -2 \\ 2 & -3 \end{bmatrix}, \& \begin{bmatrix} 3 & -2 \\ 4 & -3 \end{bmatrix}.$$

The determinant of the first one of these is not zero; therefore, the rank of A is 2. For A to have a rank of 2 it is sufficient that the determinant of <u>any one</u> of these submatrices be non-zero.

In solving a system of equations for the unknowns, it is necessary that the rank of the coefficient matrix be the same as the rank of the augmented matrix. The augmented matrix of the system is the coefficient matrix with the constant matrix, B of (1.2), appended as an added column. Consider the system

$$\begin{aligned} x_1 - 2x_2 + x_3 - x_4 &= -1 \\ 3x_1 \quad -2x_3 + 3x_4 &= 4 \\ 5x_1 - 4x_2 \quad + x_4 &= 2 \end{aligned} \tag{8.1}$$

The coefficient matrix (A) is

$$\begin{bmatrix} 1 & -2 & 1 & -1 \\ 3 & 0 & -2 & 3 \\ 5 & -4 & 0 & 1 \end{bmatrix} \tag{8.2}$$

and the augmented matrix is

$$\begin{bmatrix} 1 & -2 & 1 & -1 & -1 \\ 3 & 0 & -2 & 3 & 4 \\ 5 & -4 & 0 & 1 & 2 \end{bmatrix}. \tag{8.3}$$

The rank of both of these matrices is 2; therefore, one can solve for 2 of the unknowns in terms of the other 2 unknowns, i.e. if a value is selected for each of 2 of the unknowns, the remaining 2 are uniquely determined. Since the rank is 2 rather than 3, it is known that of the 3 equations only 2 are independent. The third can be formed from the other two. In this example the third equation can be formed by adding 2 times the first equation to the second equation. This system can be solved for the two unknowns by solving any two equations for these unknowns in terms of the remaining two. Solve the first two equations for x_1 and x_2 in terms of x_3 and x_4. From equation 2 of the system (8.1),

$$x_1 = \frac{4}{3} + \frac{2}{3}x_3 - x_4$$

Substituting this in the first equation of (8.1),

$$x_2 = \frac{7}{6} + \frac{5}{6}x_3 - x_4$$

Now, if x_3 and x_4 take on specified values, x_1 and x_2 are uniquely determined; but, there is a different value of x_1 and x_2 for each different set of values given to x_3 and x_4.

IX. Equivalence of Matrices

The augmented matrix of a system of equations can have 3 elementary transformations applied to its <u>rows</u> which will not affect the solution of the system of equations. For example, the three equations of (8.1), whose augmented matrix is (8.3), could be modified by certain operations which would not change the solution to the system. These three elementary <u>row</u> transformations are:

A. the interchange of any two rows of the augmented matrix i.e., interchanging the position of any two equations.

B. the multiplication of all elements of any row by the same <u>non-zero</u> constant

C. the addition to (subtraction from) any row of an arbitrary multiple of any other row.

Note that these transformations apply only to rows of the augmented matrix (equations of the system) and not to columns of the matrix.

Two or more matrices are said to be <u>equivalent</u> if any one matrix can be obtained from any one of the other matrices by applying these three elementary row transformations. Equivalent matrices are necessarily of the same order and rank. These transformations are those used in the Doolittle method of solving the normal equations in a least squares estimation problem.

Example:

	Equation system form	Augmented matrix form
(9.1)	$x + 2y = 3$ $-3x + 4y = 7$	$\begin{bmatrix} 1 & 2 & 3 \\ -3 & 4 & 7 \end{bmatrix}$
	is equivalent to	is equivalent to
(9.2)	$-3x + 4y = 7$ $x + 2y = 3$	$\begin{bmatrix} -3 & 4 & 7 \\ 1 & 2 & 3 \end{bmatrix}$
	is equivalent to	is equivalent to
(9.3)	$2x + 4y = 6$ $-3x + 4y = 7$	$\begin{bmatrix} 2 & 4 & 6 \\ -3 & 4 & 7 \end{bmatrix}$
	is equivalent to	is equivalent to
(9.4)	$x + 2y = 3$ $2x + 14y = 22$	$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 14 & 22 \end{bmatrix}$

The solution for x and y to each of these sets (9.1-9.4) would be the same.

(9.2) is formed from (9.1) by interchanging the positions of the two equations (rows). (9.3) is formed from (9.1) by multiplying each constant of the first equation (row) by the non-zero constant, two. (9.4) is formed from (9.1) by adding to the second equation (row) five times the first equation (row). The first equation (row) is unaffected by this transformation. Using this concept of equivalence, one can find the solution to system (9.1).

Each successive set is equivalent to every other set in this example.

	Equation system form	Augmented matrix form
(9.1)	$x + 2y = 3$ $-3x + 4y = 7$	$\begin{bmatrix} 1 & 2 & 3 \\ -3 & 4 & 7 \end{bmatrix}$
	Add 3 times equation 1 to equation 2.	
(9.5)	$x + 2y = 3$ $0x + 10y = 16$	$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 10 & 16 \end{bmatrix}$
	Divide equation 2 by 10, i.e. multiply by $\frac{1}{10}$	
(9.6)	$x + 2y = 3$ $y = 1.6$	$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 1.6 \end{bmatrix}$

Subtract 2 times equation 2 from equation 1 (or add - 2 times equation 2 to equation 1).

(9.7)	$x + 0y = -.2$ $y = 1.6$	$\begin{bmatrix} 1 & 0 & -.2 \\ 0 & 1 & 1.6 \end{bmatrix}$

Thus, the solution is $x = -.2$
$y = 1.6$

Note that in the augmented matrix the solution is the elements of the last column, when the coefficient portion of the matrix (first two columns) is in the equivalent indentity matrix form.

Example. Find the solution of the following system of 4 equations in four unknowns.

$$x_1 + x_2 + x_3 + x_4 = 1$$
$$x_1 + x_2 + x_3 - x_4 = 2$$
$$x_1 + x_2 - x_3 - x_4 = 3$$
$$x_1 - x_2 - x_3 - x_4 = 4$$

The coefficient matrix and augmented matrix are each of rank 4; therefore there is a unique value for each of the four unknowns which will satisfy each of the four equations. Writing only the succeeding equivalent forms of the augmented matrix (A) the solution is found as follows:

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 2 \\ 1 & 1 & -1 & -1 & 3 \\ 1 & -1 & -1 & -1 & 4 \end{bmatrix}$$ is equivalent to ($\sim$),

subtracting row 1 from every other row,

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 & 1 \\ 0 & 0 & -2 & -2 & 2 \\ 0 & -2 & -2 & -2 & 3 \end{bmatrix} \sim,$$

interchanging rows 2 and 4,

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & -2 & -2 & -2 & 3 \\ 0 & 0 & -2 & -2 & 2 \\ 0 & 0 & 0 & -2 & 1 \end{bmatrix} \sim,$$

dividing rows 2, 3, and 4 by (-2),

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & -1.5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -.5 \end{bmatrix} \sim,$$

subtracting row 2 from row 1,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2.5 \\ 0 & 1 & 1 & 1 & -1.5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -.5 \end{bmatrix} \sim,$$

subtracting row 3 from row 2,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2.5 \\ 0 & 1 & 0 & 0 & -.5 \\ 0 & 0 & 1 & 1 & -1 \\ 0 & 0 & 0 & 1 & -.5 \end{bmatrix} \sim$$

subtracting row 4 from row 3,

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 2.5 \\ 0 & 1 & 0 & 0 & -.5 \\ 0 & 0 & 1 & 0 & -.5 \\ 0 & 0 & 0 & 1 & -.5 \end{bmatrix} \sim$$

The coefficient part of the final equivalent matrix is an identity matrix, thus the solution for the 4 unknowns are the elements of the last column.

$$x_1 = 2.5$$
$$x_2 = -.5$$
$$x_3 = -.5$$
$$x_4 = -.5$$

X. The characteristic equation of a matrix.

The limited information single equation (LISE) method of estimating structural coefficients requires that a matrix equation of the form

(10.1) $$AX = \lambda X$$

be solved. A is a square matrix, composed of specified sums of squares and cross products of the observed data, with order equal to the number of endogenous variables in the equation and X is a vector whose elements are the structural coefficients to be estimated. λ is a set of scalars (constants), each one giving a different X. If A is nxn then there are n possible λ's which satisfy (10.1) and therefore n sets of values for X which are non-trivial, ie. for which every element of X is not zero. The LISE method provides the criterion for selecting which one of the n λ's and its associated X represents the correct estimate of the structural coefficients. This section is concerned only with finding each of the λ's (characteristic roots) and each of the associated X's (characteristic vector).

One solution to (10.1) is that X is a zero vector, ie. each $x_i = 0$. Our interest lies only in the non-trivial solutions. For what values of λ, satisfying (10.1), are there non-trivial solutions for X and what are these solutions?

(10.1) may be written

(10.2) $$AX - \lambda X = O \quad \text{or}$$
(10.3) $$(A - \lambda I)\, X = O$$

where A and I are of order n. (10.3) is satisfied if $X = O$ or if

(10.4) $$\det(A - \lambda I) = O$$

(10.5) $$A - \lambda I = \begin{bmatrix} a_{11} - \lambda & a_{12} & a_{13} \cdots\cdots & a_{1n} \\ a_{21} & a_{22} - \lambda & a_{23} \cdots\cdots & a_{2n} \\ a_{31} & a_{32} & a_{33} - \lambda \cdots & a_{3n} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} & a_{n3} \cdots & a_{nn} - \lambda \end{bmatrix}$$

$\det(A - \lambda I) = O$ is a polynomial of degree n in λ and can be solved for the n values of λ. Substituting any one of these characteristic roots in (10.3), a system of n equations in n unknowns is formed. Since $\det(A - \lambda I) = O$ the rank of the system is at most $n - 1$, therefore $n - 1$ of the structural coefficients (unknowns) can be estimated in terms of the n'th unknown. (This is the process of normalization).

Example.

$$AX = \lambda X$$

where

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}, \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ and can be thought of as the estimates of the structural coefficients of some economic structure. The problem is to find the two characteristic roots (λ_1 and λ_2) and the two associated characteristic vectors, X_1 and X_2, respectively.

$$A - \lambda I = \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{bmatrix}$$

For a non-trivial X,

$\det(A - \lambda I) = 0$, or

$$\det \begin{bmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{bmatrix} = \lambda^2 - 2\lambda - 3 = 0$$

The roots of the polynomial $\lambda^2 - 2\lambda - 3 = 0$

can be found using the quadratic formula and are $\lambda_1 = 3$
$\lambda_2 = -1$

For $\lambda_1 = 3$

$$A - \lambda I = \begin{bmatrix} 1-3 & 2 \\ 2 & 1-3 \end{bmatrix} = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$$

Note that $\det \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} = O$

Substituting this in (10. 3)

$$(A - \lambda I)\ X = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \quad X = O$$

$$\begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix} \quad \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = O$$

This gives a system of two equations in two unknowns.

$$-2x_1 + 2x_2 = O$$

$$2x_1 - 2x_2 = O$$

Note that equation two is the negative of equation 1; therefore there is only one independent equation leading to the solution

$$x_1 = x_2$$

Thus the solution for $\lambda_1 = 3$ is any set of values for which $x_1 = x_2$.

The characteristic equation

$$(10.6) \qquad \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

is satisfied by $\lambda_1 = 3$ and any set of values, x_1 and x_2, which are equal.

For the second characteristic root $\lambda_2 = -1$, the solution of the characteristic vector is

$$x_1 = -x_2 \ .$$

Thus $\lambda_2 = -1$ and any set of values, x_1 and x_2, which are equal in magnitude but different in sign is a solution of (10. 6).

XI. An example of matrix algebra in multiple regression problems.

Using matrix notation let us develop a method of estimating the coefficients of the independent variables (x_1, x_2, $\cdots$, x_5) in the following regression relation.

(11.1) $$y_t = B_1 x_{1t} + B_2 x_{2t} + B_3 x_{3t} + B_4 x_{4t} + B_5 x_{5t} + u_t$$

All variables are measured as deviations from the means of observed data of each specific variable; thus, the intercept ("a" value) is zero. y_t, x_{1t}, x_{2t} $\cdots$ x_{5t} are the observed deviations for the t'th time period. u_t is the random error for the t'th time period and represents the effect of other explanatory variables on y, the dependent variable. B_1, B_2, $\cdots$ B_5 are the structural parameters to be estimated by least squares. Let T be the number of observations. If b_1, b_2, $\cdots$, b_5 are the least squares estimates of B_1, B_2, $\cdots$ B_5, respectively, the estimating equation for y is

(11.2) $$y_t = b_1 x_{1t} + b_2 x_{2t} + b_3 x_{3t} + b_4 x_{4t} + b_5 x_{5t} .$$

Let

(11.3) $$Y = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ y_T \end{bmatrix}$$ be the vector of observed values of the dependent variable;

(11.4) $$X = \begin{bmatrix} x_{11} & x_{21} & x_{31} & x_{41} & x_{51} \\ x_{12} & x_{22} & x_{32} & x_{42} & x_{52} \\ x_{13} & x_{23} & x_{33} & x_{43} & x_{53} \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_{1T} & x_{2T} & x_{3T} & x_{4T} & x_{5T} \end{bmatrix}$$

be the matrix of observed values of the 5 independent variables;

(11.5) $$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \end{bmatrix}$$ be the vector of estimated coefficients;

(11.6) $$U = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ \cdot \\ \cdot \\ \cdot \\ u_T \end{bmatrix}$$ be the vector of unobserved error terms;

and

(11.7) $$B = \begin{bmatrix} B_1 \\ B_2 \\ B_3 \\ B_4 \\ B_5 \end{bmatrix}$$ be the vector of parameters to be estimated.

Then

(11.8) $$Y = XB + U$$

is the set of T observation equations

(11.9)
$$\begin{aligned} y_1 &= B_1 x_{11} + B_2x_{21} + B_3x_{31} + B_4x_{41} + B_5 x_{51} + u_1 \\ y_2 &= B_1 x_{12} + B_2x_{22} + B_3x_{32} + B_4x_{42} + B_5x_{52} + u_2 \\ &\vdots \\ y_T &= B_1 x_{1T} + B_2x_{2T} + B_3x_{3T} + B_4x_{4T} + B_5 x_{5T} + u_T \end{aligned}$$

The least squares method finds the estimates of B, denoted b, which will minimized the sums of the squares of u_t, i.e. the estimates which minimize

$$U'U$$

Solving (11.8) for B, we get

$$Y = XB + U \; ;$$

(11.8) pre-multiply by X',

(11.9) $$X'Y = (X'X)\,B + X'U\,.$$

(X'X) is the matrix of sums of squares and cross products of the deviations of the independent variables.

(11.10) $$(X'X) = \begin{bmatrix} \sum_t x_{1t}^2 & \sum_t x_{1t}x_{2t} & \sum_t x_{1t}x_{3t} & \sum_t x_{1t}x_{4t} & \sum_t x_{1t}x_{5t} \\ \sum_t x_{1t}x_{2t} & \sum_t x_{2t}^2 & \sum_t x_{2t}x_{3t} & \sum_t x_{2t}x_{4t} & \sum_t x_{2t}x_{5t} \\ \sum_t x_{1t}x_{3t} & \sum_t x_{2t}x_{3t} & \sum_t x_{3t}^2 & \sum_t x_{3t}x_{4t} & \sum_t x_{3t}x_{5t} \\ \sum_t x_{1t}x_{4t} & \sum_t x_{2t}x_{4t} & \sum_t x_{3t}x_{4t} & \sum_t x_{4t}^2 & \sum_t x_{4t}x_{5t} \\ \sum_t x_{1t}x_{5t} & \sum_t x_{2t}x_{5t} & \sum_t x_{3t}x_{5t} & \sum_t x_{4t}x_{5t} & \sum_t x_{5t}^2 \end{bmatrix}$$

Note that (X'X) is symmetrical, thus its inverse will also be symmetrical. (X'Y) is the matrix (column vector) of sums of cross products of the dependent variable with each of the independent variables.

(11.11) $$(X'Y) = \begin{bmatrix} \sum_t x_{1t}y \\ \sum_t x_{2t}y \\ \sum_t x_{3t}y \\ \sum_t x_{4t}y \\ \sum_t x_{5t}y \end{bmatrix}$$

Note that

(11.12) $$X'Y = (X'X)\,b$$

is the set of normal equations for the regression relation (11.1). Pre-multiplying (11.12) by the inverse of (X'X), we obtain

$$(X'X)^{-1}\ (X'Y) = (X'X)^{-1}\ (X'X)b, \text{ or}$$

(11.13) $$(X'X)^{-1}\ (X'Y) = Ib = b.$$

(11.13) gives the least squares estimates of the structural parameters of (11.1).